Validity and reliability of the Japanese version of the Therapeutic Factors Inventory-19 (TFI-19J): A study of family peer education self-help groups

Validity and reliability of the Japanese version of the Therapeutic Factors Inventory-19 (TFI-19J) - A study of family peer education self-help groups

Japan Journal of Nursing Science (2016) **13** 135–146

doi:10.1111/jjns.12098

Running head:

THERAPEUTIC FACTORS INVENTORY-19

Masako KAGEYAMA,[1] Yukako NAKAMURA,[2] Sayaka KOBAYASHI[3] and Keiko YOKOYAMA[4]

[1]*Department of Community Health Nursing, Graduate School of Medicine, The University of Tokyo, Tokyo,* [2]*Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya,* [3]*National Center of Neurology and Psychiatry, National Institute of Mental Health, Kodaira and* [4]*Department of Nursing, Faculty of Health Sciences, Saitama Prefectural University, Koshigaya, Japan*

Correspondence: Masako Kageyama, Department of Community Health Nursing, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email: kageyama-tky@umin.ac.jp

Received 10 October 2014; accepted 21 July 2015.

Abstract

Aim: Therapeutic factors are crucial mechanisms that promote change in self-help group members. Measuring therapeutic factors may improve practitioners' skills for assessment in whole-group contexts. We, therefore, examined the validity and reliability of a Japanese version of the Therapeutic Factors Inventory-19.

Methods: The Therapeutic Factors Inventory-19 was examined using a self-report questionnaire completed by members of 38 family peer education self-help groups. The instrument measured the following four factors: instillation of hope, secure emotional expression, awareness of relational impact, and social learning.

Results: Participants were 246 group members. Test–retest reliability was analyzed using data from 46 participants. Confirmatory factor analysis showed a GFI of 0.85 and an RMSEA of 0.0088. Multitrait scaling analyses showed that items for instillation of hope and secure emotional expression factors correlated higher with their own factors than other factors. Each factor and the total average of the 19 items were significantly correlated with the Group Benefit Scale and Client Satisfaction Questionnaire. When level of interaction with other members was higher, subjects perceived a stronger presence of therapeutic factors. The intraclass correlation coefficients of each factor at a week interval were 0.848–0.915. The Cronbach's alpha of each factor and all items ranged from 0.767 to 0.960.

Conclusions: In the case of family peer education self-help groups, there is acceptable validity and reliability for the average score of all items, and for the instillation of hope and secure emotional expression factors. However, more work is needed to increase the generalizability.

Keywords: group processes, group psychotherapy, process assessment, reliability

and validity, self-help groups

Introduction

As part of caring for people with various difficulties, nurses have the opportunity to facilitate and/or support groups. In Japan, nurses working in clinical and community settings are often involved with therapy, support, and self-help groups (Japanese National Federation of Families of the Mentally Ill, 1998; Suzuki, 1999; Tanimoto & Kitaoka-Higashiguchi, 2005). In order to appropriately facilitate and/or support groups, nurses first need to assess and promote group-as-a-whole processes (Kageyama, 2011). However, there are few existing Japanese scales to assess the processes used in a group context. One way to assess group processes is using the following therapeutic factors proposed by Yalom (1995): (1) instillation of hope, (2) universality, (3) imparting of information, (4) altruism, (5) the corrective recapitulation of the primary family group, (6) development of socializing techniques, (7) imitative behavior, (8) interpersonal learning, (9) group cohesion, (10) catharsis, and (11) existential factors. These therapeutic factors have a rich history in the literature on therapy groups, and have been described as crucial mechanisms that promote change in group therapy (Joyce, MacNair-Semands, Tasca, & Ogrodniczuk, 2011). Yalom (1995) created the therapeutic factors so that these were intrinsic to the therapeutic process not only in therapy groups but also in support groups and self-help groups. The focus of therapeutic groups is on change or growth, rather than providing a cure (Yalom, 1995). Therefore, the therapeutic factors are manifested in support and self-help groups, as well as in therapy groups (Kurtz, 1997). Self-help groups are voluntary, small group structures for mutual aid and the accomplishment of a special purpose (Katz & Bender, 1976). It was acknowledged that support provided by practitioners for self-help groups, including consultation, co-leadership, and referral, is essential; however, this type of

involvement is a new area of nursing practice and research (Adamsen & Rasmussen, 2001; Carlsen, 2003). The Japanese government recently announced a community health policy to promote the activities of self-help groups (Ministry of Health, Labour and Welfare, 2012). Under this policy, practitioners are encouraged to provide support for self-help groups by gaining appropriate skills. Practitioners need to assess the whole group in order to plan effective strategies; however, assessment is one of most challenging aspects of group practice (Toseland & Rivas, 2009). Measures to assess therapeutic factors may improve practitioners' skills for whole-group assessment and support them in developing a tailored manner (Lese & MacNair-Semands, 2000).

A number of measures have been developed to assess the above therapeutic factors, including objective measures (Hastings-Vertino, Getty, & Wooldridge, 1996) and subjective measures of group members' perspectives (Fuhriman, Drescher, Hanson, Henrie, & Rybicki, 1986; Joyce *et al*., 2011; Lese & MacNair-Semands, 2000; Lieberman, Yalom, & Miles, 1973; MacNair-Semands, Ogrodniczuk, & Joyce, 2010; Stone, Lewis, & Beck, 1994). Subjective and easy-to-use measures may have a wider range of utilization, even for less experienced practitioners. Further, the Therapeutic Factors Inventory-19 (TFI-19; Joyce *et al*., 2011) may be one of the simplest of these self-report measures, given that it is limited to 19 items. However, there is no Japanese version of any instrument that specifically measures Yalom's therapeutic factors. If practitioners can use a Japanese version of the established English TFI-19, they may be able to assess and support self-help groups in a more appropriate manner. Our aims in this study were to develop a Japanese version of the TFI-19 (TFI-19J) and assess its validity and reliability.

Methods

General Descriptions of the TFI-19

Originally, the TFI-19 was developed based on the 99-item Therapeutic Factors Inventory (TFI) that assesses 9 items for each of the 11 therapeutic factors (Lese & MacNair-Semands, 2000; MacNair-Semands & Lese, 2000). Each therapeutic factor is used as a subscale. A limitation to the widespread use of the TFI is its length. As a result, a 23-item short version known as the TFI-Short Form (TFI-S) was developed (MacNair-Semend *et al*., 2010). Subsequently, the TFI-19 was created to further refine the constituent items from TFI-S (MacNair-Semend *et al*., 2010). Yalom & Leszcz (2005) acknowledge that the 11 therapeutic factors do not function independently. Consequently, in the TFI-19, Joyce *et al*. (2011) used the following four combined factors derived from the 11 original therapeutic factors: instillation of hope (IH; 4 items), secure emotional expression (SEE; 7 items), awareness of relational impact (ARI; 5 items), and social learning (SL; 3 items). In the TFI-19, IH includes the two therapeutic factors of hope and universality. "The recognition of universality among the members promotes hope" (Joyce *et al*., 2011, p.3). SEE reflects a sense of connection with other group members that include concepts of group cohesion, self-disclosure, and catharsis. The factor appears to reflect one's sense of safety and comfort in a group; thus, "it may be associated with the members communicating openly and honestly" (Joyce *et al*., 2011, p.3). ARI refers to "a connection between interpersonal experiences and cognitive-affective factors associated with gaining insight; this factor would be expected to increase over time as members receive feedback from others in the group" (Joyce *et al*., 2011, p.3). Finally, SL emphasizes skills acquired through behavioral processes (Joyce *et al*., 2011; MacNair-Semands *et al*., 2010). "The transition from

insight to action is considered the most difficult to accomplish in therapy" groups (Joyce *et al*., 2011, p.3).

The TFI-19 is a self-report measure that assesses individual group members' perceptions of the presence of the four global therapeutic factors rather than how effective the factors are in the group context using the following statement: "Please rate the following statements as they apply to your experiences in your group by circling the corresponding number, using the following scale." Each item is rated on a Likert-type scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Each score on the four factors can be weighted or calculated as a simple average of associated items. For the purposes of the current study, we chose to use the simple average score in our analysis.

Development of the TFI-19J

The TFI-19 was developed in groups led by therapists (Joyce *et al*., 2011), and we tested it in self-help groups led by lay-trained peers. The TFI-19 is an instrument used to assess the group status in regard to how therapeutic factors function, not to assess the way to conduct the intervention itself. Therefore, we considered that the TFI-19 should work well in self-help groups. The TFI-19J was developed as follows: To begin, the first author contacted the developer of the TFI-19 and obtained permission to develop a Japanese version. Next, each author, including registered nurses who are also teaching staff in psychiatric and public health nursing, a licensed clinical psychologist, and a licensed psychiatric social worker, translated the English version of the TFI-19 into Japanese. We all have rich backgrounds in supporting self-help groups and facilitating support groups, and two of us have in the past facilitated therapy groups. Two have acted as practical instructors of therapy groups and the other two have written nursing textbooks about these groups. Together, the researchers discussed the process

and reached a consensus about a Japanese translation, deciding to use natural Japanese so that the new version may be used in self-help groups, support groups, and therapy groups. We also discussed whether the translations were equivalent in terms of the meaning of the original items and theoretical construction, and whether the translation was suitable for use in a Japanese cultural background. Japanese people are unlikely to express negative emotions to others, so item 5, i.e., "It's okay for me to be angry in group," was the most difficult to translate. The literal translation may be too emotional and, therefore, distort the meaning of the factor; further, it may cause a possible floor effect, so we translated the item as "It's okay for me to express my anger in group." Then, 22 families who belonged to self-help groups were asked to respond to the items and provide feedback about whether or not they were understandable. We also checked for extreme response distribution, especially for item 5. After items were changed based on the opinions of the 22 families, six additional families, each with rich experience attending support and self-help groups for families of persons with mental disorders, were asked to respond to the version and give feedback about whether it was understandable and appropriate for use in the group context. Subsequently, the Japanese version was revised until the respondents found it easy to understand; further, we confirmed the semantic and conceptual equivalence of the translated instrument with the original version. The Japanese translation was then reverse-translated by a native English speaker. Finally, the contact person for the TFI-19 checked the correctness of the reverse translation. The Japanese translation was modified several times until agreement was reached between the families and the contact person.

Survey for Testing the Validity and Reliability of the TFI-19J

We tested the validity and reliability of the TFI-19J in family peer education

self-help groups. In this study, the groups were facilitated by three to six host members that belonged to local family self-help organizations. The groups also included guest members, who were family members with or without membership in family self-help organizations. The groups were small (included less than 15 participants), and had closed membership. In addition, each group met for a series of 5 sessions (3 hours per session). The combination of textbook readings and experience sharing was repeated at each meeting.

A self-report questionnaire survey was administered to members of 38 self-help groups that had their final session between October 2013 and March 2014. In the last session, a representative host member from each group distributed a questionnaire to each host and guest member. Each member answered the questionnaire individually at home following the session and sent it directly back to the researcher. To test the test-retest reliability of the TFI-19J, a second copy of the questionnaire was provided to host and guest members in 10 of the groups. These participants were asked to respond to and return the TFI-19J a second time one week after they completed the first survey.

Data Analysis

First, subjects for analysis were determined after removing respondents who did not fully complete the TFI-19J or who responded incorrectly to the following item: “No family members of persons with mental illness joined the group.” The respondents who selected 6 or 7 (strongly agree) in response to the false item were excluded from analysis. The criterion for joining group members was that they were family members of persons with mental illness so that the false item did not reflect real groups. We used this false item in order to test the TFI-19J as accurately as possible, since it allowed us to select respondents who answered the questions appropriately. Using such a method is

recommended in scale development to avoid response distortion resulting from intentional or unintentional incorrect answers (Murakami, 2006). We have a lot of experience with conducting surveys of family members, some of whom have poor concentration. One of the reasons for this is that they have little capacity in terms of time and mental resources because they live with the person with mental illness and provide care on a daily basis (Chiba Prefecture Family Association of Persons with Mental Disorders, 2009; National Association of Family Groups on Mental Disorders, 2006). This is why the false item was used.

Next, we checked the basic score distribution in order to assess floor and ceiling effects. These effects are considered to be present if the mean plus standard deviation (SD) > the highest possible score or mean minus SD < the lowest possible score. Correlations between the four factors, the number of factors, confirmatory factor analysis, and a multitrait scaling analysis were used to check construct validity. The number of factors was tested by eigenvalues and the scree test. According to the Kaiser-Guttman rule, the number of factors were determined when an eigenvalue > 1.0. The scree test plotted eigenvalues and determined where factors leveled off. The confirmatory factor analysis was conducted based on the model hypothesized in the original TFI-19 (Joyce *et al*., 2011) that reflected correlations between the four factors and first-order latent factors. In addition, the goodness-of-fit of the hypothesized model were assessed using a number of tests, including: χ^2 (smaller value equals better fit), GFI (the goodness-of-fit index; ideally >0.9), and RMSEA (the root mean square error of approximation; ideally <0.08). For the multitrait scaling analysis, convergent validity was analyzed using the Spearman correlation at ≥0.4 (Fayers & Machin, 2000). Discriminant validity was supported when each item had a higher correlation with its

own factor (corrected for overlap) than with other factors. Scaling success was counted when the item to own-factor correlation was higher than the correlations of the item to other factors.

The concurrent validity of the original TFI-19 was tested using the Group Climate Questionnaires-Short Form (GCQ-S; MacKenzie, 1983). It is a self-report measure designed to assess individual members' perceptions of a group's therapeutic environment, and a sample item is as follows: "The members like and care about each other." However, no Japanese version of the GCQ-S or any other scale to directly assess group therapeutic environment exists. As a result, we used the Japanese versions of scales that are not equivalent to GCQ-S but reflect the group process to assess concurrent validity. Specifically, we examined correlations with the Group Benefit Scale (Maton, 1988) and Client Satisfaction Questionnaire (CSQ-8) (Larsen, Attkisson, Hargreaves, & Nguyen, 1979; Tachimori & Ito, 1999). The Group Benefit Scale is a 5-item self-report scale designed to assess individual members' appraisals of the personal benefits received from group involvement. Each item ranges from 1 (not at all accurate) to 5 (completely accurate), with higher scores indicating higher appraisal. The Japanese version of the Group Benefit Scale was developed and used among family group members that were similar to the subjects of the current study (Kageyama & Oshima, 2007). The CSQ-8 is an 8-item self-report scale that measures consumer satisfaction of health services. Each item can be rated from 1 to 4, with higher scores indicating higher satisfaction. The CSQ-8 was used only for guest members. We also confirmed concurrent validity by examining the relationship between the four factors and level of length and frequency of interaction with other members. Since the factors are expected to increase over time in the group because of accumulating interaction opportunities

(Joyce *et al*., 2011), we hypothesized that the factors would be higher with a higher level of interaction with other members. The level of interaction with other members was divided into the following three levels: low, moderate, and high. The subjects with a low level were guest members without membership in a self-help organization. Thus, they only knew other members in 5 sessions through group participations. The subjects with moderate level of interaction were guest members with existing organization memberships, signifying that they knew other group members prior to their participation in the group and typically met once a month in addition to attending the five sessions. Finally, host members were classified as subjects with a high level of interaction because they knew other group members prior to their involvement in the group, and because they met frequently—around three times a month—to prepare for the group meetings, in addition to attending the five sessions. Since grouped data is nested within groups, we used a generalized linear mixed model with groups as a random effect, which took into account the extra component of variation due to the nested design (Donner, Brown, & Brasher, 1989).

Test-retest reliability was analyzed using the intraclass correlation coefficients (ICC) of each factor over a one-week period. An ICC of 0.7 is often recommended as the minimum standard for reliability (Terwee *et al*., 2007). Internal consistency reliabilities were checked using Cronbach's alpha.

All analyses were conducted using SAS version 9.4 (SAS, North Carolina, United Sates), with the exception of ICCs, which were analyzed using SPSS Statistics version 20 (IBM, SPSS Statistics for Windows, New York, United States).

Ethical Considerations

This study was approved by the Ethics Committee at the University of Tokyo,

Faculty of Medicine, at the first author's University (No.10146; May 14, 2013). All subjects were informed of the study's purpose and their right to refuse participation. Participants were considered to have provided consent if they returned a questionnaire to the researchers.

Results

Subject Demographics

Initially, questionnaires were distributed to 224 guest members and 164 host members of 38 self-help groups during the final session, between October 2013 and March 2014. Questionnaires were returned by a total of 296 members (163 guest members and 133 host members). Of the 296 questionnaires that were returned, 26 had one or more missing item(s) on the TFI-19J. In addition, in 22 of the remaining 270 questionnaires, members selected incorrect ratings in response to the false item (i.e., a 6 or 7). Therefore, a total of 246 members (137 guest members and 111 host members) comprised the subjects for analysis.

The second questionnaire for studying test-retest reliability was provided to 73 members from 10 groups. Of these 73 participants, 57 members returned the questionnaires and 46 completed the 19-items and correctly responded to the false item. Thus, there were a total of 46 subjects included in the analysis of test-retest reliability.

Participants' demographic characteristics are shown in Table 1. The subjects had an average age of 65.4, approximately 69% were female, and 79% had the membership in a local family self-help organization. The majority of family members who attended the group had male children aged 30–40 with schizophrenia.

[Insert Table 1]

Score Distributions

The score distribution of each item is shown in Table 2. To compare the floor and ceiling rates to the original TFI, the table presents the mean, SD, mean-SD, and mean+SD of the TFI-S (MacNair-Semend *et al*., 2010). The means for the majority of items were 5 five or higher, except for item number 5: "It's okay for me to be angry in group." The items where the mean+SD >7.0 were numbers 2, 3, 9, 14, 4, 18, and 1. However, over half of the subjects selected the highest score only for item number 9. There were no items where over 60% of participants chose the highest or lowest scores.

[Insert Table 2]

Construct Validity

As shown in Table 3, all Spearman correlation coefficients between factors were high (0.781–0.949). Only one factor had an eigenvalue >1.0 (10.820). The eigenvalue of the second factor was 0.506, and the scree test demonstrates where it levels off. The results of confirmatory factor analysis are showed in Table 4. Here, the χ^2 was high, with a GFI of 0.85, indicating that the data was not a good fit. In addition, the RMSEA was 0.088, which is not ideal. Consequently, the GFI indices resulted in numbers that were inadequate and slightly worse than original TFI-19 data (see Table 4). The results of multitrait scaling analyses are shown in Table 5. In this case, the correlations of all items with their own factors were high. In addition, the IH and SEE items were correlated slightly higher with their own factors than with other factors. The scaling successes were represented in the findings of 100.0% in IH, 95.5% in SEE, 60.0% in ARI, and 0.0% in SL.

[Insert Table 3]

[Insert Table 4]

[Insert Table 5]

Concurrent Validity

As shown in Table 6, each factor and the total average of the 19 items were significantly related to the Group Benefit Scale and CSQ-8. Thus, the subjects perceived the presence of therapeutic factors as stronger, appraised higher benefits from group involvement, and gained higher satisfaction from the group. As presented in Table 7, the level of interaction with other members was significantly related to all factors and the total average. Furthermore, subjects with a higher level of interaction with other members perceived a stronger presence of the therapeutic factors.

[Insert Table 6]

[Insert Table 7]

Reliability

The ICC of each factor between the two times the test was completed ranged from 0.848 to 0.915. The Cronbach's alpha of each factor and all items ranged from 0.767 to 0.960.

Discussion

Validity and Reliability of TFI-19J

In the current study, we developed the TFI-19J and tested its validity and reliability. The results of the ICC and Cronbach's alpha showed appropriate levels of reliability for the TFI-19J.

According to the validity testing, the score distribution of seven items showed the presence of ceiling effects. Other than these items, the scores for most items were rated over 5. As a whole, the scores were higher than those found in the analysis of the

TFI-S (MacNair-Semend *et al*., 2010). One possible reason for this difference relates to the subjects' characteristics; specifically, being family members of persons with mental illness, rather than individuals who themselves had been diagnosed with mental disorders. In contrast to our study findings, individuals involved in the TFI-S study were persons who had mental illness, 70% of whom suffered from depression (MacNair-Semend *et al*., 2010). Thus, a possible explanation is that persons with depression may score lower than individuals without a mental illness. Another possible reason is the length of time that the group participants knew each other. Specifically, the therapy groups assessed in the TFI-S study were part of an 18-week, time-limited, and intensive treatment group. In contrast, the subjects of this study included people with memberships in the self-help organization and, as a result, many had known each other for a more substantial amount of time (e.g., some knew each other for over a decade). Accordingly, individuals with a higher level of interaction with other members had scores that were significantly higher than guest members without an organizational membership (see Table 7). Consequently, the length of time that group members know each other might facilitate higher scores. In addition, with the exception of one item, less half of the participants chose the highest score on the majority of items. On the basis of the above discussion, there is not enough evidence to delete or modify existing items. That said, further studies with subjects from other types of treatment groups in Japan are needed.

Next, the results of testing the concurrent validity of the TFI-19J showed that the four factors and the total average score significantly correlated with the Group Benefit Scale and CSQ-8, as well as the level of interaction with other members. Although we could not use the GCQ-S, the results signify a moderate level of concurrent validity.

The analysis of construct validity revealed there were high correlations among factors, with only one factor having an eigenvalue > 1.0. There were substantial differences in eigenvalues for the other factors. The non-ideal results of the confirmatory factor analysis signify the inadequacy of the four separate factors and the single factor structures. In addition, the results of multitrait scaling analysis showed that only IH and SEE factors were marginally successful. On the basis of these results, the total average score, and the IH and SEE factors can be considered to have an acceptable level of construct validity, while the ARI and SL factors do not. One possible reason for the factor structures found in this study being different from the original theoretical construct is the effect of cultural backgrounds. The original TFI-19 was developed in the US and did not address cultural issues; however, use of the therapeutic factors has become common among Japanese practitioners. In fact, the books by Yalom (1995) and Yalom & Leszcz (2005) were translated into Japanese. Textbooks for nursing students often cited the Japanese translation about therapeutic factors without discussion of cultural differences. We do not deny that there may be an effect of cultural background; however, we did not find evidence supporting it. Therefore, the effect of cultural background needs to be discussed in future research. On the other hand, there is evidence supporting the fact that the differences in the theoretical construct are due to the differences between therapists-led groups and laypersons-led self-help groups.

IH and SEE Factors in the Context of Self-Help Groups

Yalom's therapeutic factors of the TFI-19 emerge theoretically in self-help groups. However, only the IH and SEE demonstrated construct validity, which implies that subjects only appropriately perceived these two concepts. Interestingly, the IH and SEE factors include concepts of hope, universality, group cohesion, self-disclosure, and

catharsis. Similar to the results of other researchers (e.g., Kurts, 1997; Levy, 1979), Diefenbeck, Klemm, and Hayes (2014) showed that group cohesion, catharsis, imparting information, and universality, which are included in Yalom's therapeutic factors, often emerged in self-help groups. In the current study, host members tried to facilitate a high level of cohesion and expressed empathy to create an atmosphere of hope for guest members, allowing them to feel comfortable about sharing their experiences (Kageyama & Yokoyama, 2012). On the basis of these study findings and explanations, it is likely that the concepts related to IH and SEE will consistently appear in self-help groups, such as the ones used in the current study, which explains why participants could more accurately evaluate these two factors.

Conversely, Diefenbeck et al. (2014) showed that interpersonal learning, corrective recapitulation of the primary family group, imitative behavior, and development of socializing techniques were absent or virtually non-existent in self-help groups. These findings are similar to those of Levy (1979). The four therapeutic factors may be similar or more strongly related to the concepts of ARI and SL. Possible explanations for the finding of the absence of the four factors by Diefenbeck et al. (2014) include the nature of self-help groups and the lack of a professional leader. Levy (1979) also indicated that such factors appear to be more overtly controlling and behaviorally manipulative. Specifically, these activities are rarely found in natural settings and are more characteristic of the "artificial" settings constructed by therapists. In contrast, in the current study, rather than therapists, trained host facilitators that were also family members led groups. This may be an important reason why ARI and SL did not have significant construct validity in this study.

After comparing our findings with those from previous studies, we believe that

only the IH and SEE factors from the TFI-19J are appropriate for use in the context of self-help groups. However, further research needs to be conducted in other types of therapeutic groups (i.e., therapy groups and support groups) to validate our findings by extending them to other types of groups.

Limitations and Further Research

An important limitation is that the validity and reliability were tested only within the context of self-help groups. For this reason, only all items, and the IH and SEE factors showed acceptable validity. The findings show that only the IH and SEE factors may be used as subscales in self-help groups. However, further research to compare to other types of groups is needed to validate the findings. Further, we could not deny the effect of cultural background on differences between the factor structure found in this study and the original theoretical construct. The effect of cultural background needs to be discussed in future research. Another limitation is that we could not assess the concurrent validity with the GCQ-S used in developing the original TFI-19. One further limitation is that we tested the TFI-19J at a single point in time (at the end of the five sessions); however, the therapeutic factors are expected to increase over time as the groups continue (Joyce *et al*., 2011). Thus, the evaluation of the measure at several time points could increase the sensitivity of the test to change, and allow predictive validity to be established. In the current study, we did not conduct the testing a multiple time points because we did not want to burden our participants. Further research could be conducted to assess the TFI-19J over time, which will be useful to detect the effectiveness of therapeutic groups and to evaluate the effects of various strategies used by group practices.

To address the above limitations, further studies are needed to test the validity of

the TFI-19J with other types of therapeutic groups, especially in the context of therapy groups in Japan. Moreover, the concurrent validity of the TFI-19J is recommended to be tested with a Japanese version of GCQ-S, when this has been developed. Changes in TFI-19J scores over time need to be examined in the future.

Implications for Practice

The development of the TFI-19J will enable Japanese practitioners—even lay group leaders—to assess the process of self-help groups. Practitioners seeking to support self-help groups can use the TFI-19J to promote a higher quality of group practice. The average score of all items will be useful for practitioners to comprehensively assess the group process. IH and SEE are basic and important factors. If a low IH score is received, practitioners can check the composition of group members. Selection of group members is important for effective group practice (Yalom, 1995) and the high homogeneity of members can enhance their sense of universality (Kurtz, 1997). “The recognition of universality among the members promotes hope” (Joyce et al., 2011, p.3). One workable solution to address the low homogeneity of members is creating subgroups where member numbers allow for this (Kurtz, 1997). Practitioners could suggest that group leaders create subgroups of members with a high degree of commonality. If a low SEE score is received, practitioners can check whether the group environment is safe enough to allow participants to share their private experiences and feelings. Members will be unwilling to their share own private experiences unless the group follows the guidelines of confidentiality (Toseland & Rivas, 2009). If the practitioners judge that there are concerns in relation to safety, they can tell the group leader and members about the importance of following the rules of confidentiality or non-critical talk among members. High quality group practices will

contribute to members' change and growth through participating in self-help groups.

Conclusion

In conclusion, in the case of family peer education self-help groups, there is acceptable validity and reliability for the average score of all items, and for the IH and SEE factors of the TFI-19J. However, more work is needed to increase the generalizability of the TFI-19J.

Acknowledgements

The authors would like to thank all of cooperating subjects.

Disclosure and Conflict of Interests

This work was supported by JSPS KAKENHI Grant Number 23593453. The authors report no conflicts of interest.

Authors' Contributions

YN, SK and KY translated the original scale into Japanese and interpreted the data.

References

Adamsen, L., & Rasmussen, J. M. (2001). Sociological perspectives on self-help groups: Reflections on conceptualization and social processes. *Journal of Advanced Nursing*, *35*(6), 909–917. doi: 10.1046/j.1365-2648.2001.01928.x

Carlsen, B. (2003). Professional support of self-help groups: A support group project for Chronic Fatigue Syndrome patients. *British Journal of Guidance & Counseling*, *31*(3), 289–303. doi: 10.1080/0306988031000147884

Chiba Prefecture Family Association of Persons with Mental Disorders. (2009). *Needs in life and welfare of persons with mental disorders living in the community and their families 2008*. Chiba, Japan: Author (in Japanese).

Diefenbeck, C. A., Klemm, P. R., & Hayes, E. R. (2014). Emergence of Yalom's therapeutic factors in a peer-led, asynchronous, online support group for family caregivers. *Issues in Mental Health Nursing*, *35*, 21–32. doi: 10.3109/01612840.2013.836260

Donner, A., Brown, K. S., & Brasher, P. (1989). A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979–1989. *International Journal of Epidemiology*, *19*(4), 795–800.

Fayers, P. M., & Machin, D. (2000). *Quality of life: Assessment, analysis and interpretation*. New York: Wiley.

Fuhriman, A., Drescher, S., Hanson, E., Henrie, R. & Rybicki, W. (1986). Refining the measurement of curativeness – An empirical approach. *Small Group Research*, *17*, 186–201. doi: 10.1177/104649648601700204

Hastings-Vertino, K., Getty, C., & Wooldridge, P. (1996). Development of a tool to measure therapeutic factors in group process. *Archives of Psychiatric Nursing*,

10(4), 211–228.

Japanese National Federation of Families of the Mentally Ill. (1998). *Actual situation and prospects of family groups' activities from national survey of family groups' activities*. Tokyo, Japan: ZENKAREN (in Japanese).

Joyce, A. S., MacNair-Semands, R., Tasca, G. A., & Ogrodniczuk, J. S. (2011). Factor structure and validity of the Therapeutic Factors Inventory–Short Form. *Group dynamics: Theory, Research, and Practice*. Published online ahead of print. doi: 10.1037/a0024677

Kageyama, M., & Oshima, I. (2007). Intervention study for promoting partnerships between professionals and self-help groups of families of individuals with severe mental illness in Japan. *Japanese Journal of Public Health, 54*, 314–323.

Kageyama, M. (2011). Assessment for group support. In S. Murashima (Ed.), *Support skills for public health nursing* (pp. 302–310). Tokyo, Japan: Medical Friend (in Japanese).

Kageyama, M., & Yokoyama, K. (2012). Skills to facilitate a family-led education program for mental illness. *Japanese Journal of Psychiatric Rehabilitation*, *16*(1), 62–69 (in Japanese).

Katz, A., & Bender, E. (1976). *The strength in us: Self-help groups in the modern world.* New York, NY: Franklin-Watts.

Kurtz, L. F. (1997). *Self-help and support groups: A handbook for practitioners*. Thousand Oaks, CA: Sage.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, *2*, 197–207.

Lese, K. P., & MacNair-Semands, R. R. (2000). The Therapeutic Factors Inventory: Development of a scale. *Group, 24*, 303–317. doi: 10.1023/A:1026616626780

Levy, L. H. (1979). Processes and activities in groups. In M. A. Lieberman & L. D. Borman (Eds.), *Self-help groups for coping with crisis: Origins, members, processes, and impact* (pp. 234–271). San Francisco, CA: Jossey-Bass.

Lieberman, M. A., Yalom, I. D., & Miles, M. B. (1973). *Encounter groups: First facts.* New York, NY: Basic Books.

MacKenzie, K. R. (1983). The clinical application of a group measure. In R. R. Dies & K. R. MacKenzie (Eds.), *Advances in group psychotherapy: Integrating research and practice* (pp. 159–170). New York, NY: International University Press.

MacNair-Semands, R. R., & Lese, K. P. (2000). Interpersonal problems and the perception of therapeutic factors in group therapy. *Small Group Research*, *31*, 158–174.

MacNair-Semands, R. R., Ogrodniczuk, J. S., & Joyce, A. S. (2010). Structure and initial validation of a short form of the Therapeutic Factors Inventory. *International Journal of Group Psychotherapy, 60*(2), 245–281. doi: 10.1521/ijgp.2010.60.2.245

Maton, K. I. (1988). Social support, organizational characteristics, psychological well-being, and group appraisal in three self-help group populations. *American Journal of Community Psychology*, *16*(1), 53–77.

Ministry of Health, Labour and Welfare. (2012). *Modification of basic policy regarding promotion strategies for community health (notice no. 0731-8)* (in Japanese). Retrieved from: http://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-

Kenkoukyoku/0000050854.pdf

Murakami, Y. (2006). *How to develop a psychological scale*. Kyoto, Japan: Kitaohji Syobo (in Japanese).

National Association of Family Groups on Mental Disorders. (2006). *Fourth national survey report about family needs: Current living condition and attitude of persons with mental disorder and their families*. Tokyo, Japan: National Association of Family Groups on Mental Disorders (in Japanese).

Stone, M. H., Lewis, C. M., & Beck, A. P. (1994). The structure of Yalom's Curative Factor Scale. *International Journal of Group Psychotherapy, 44*(2), 239–245.

Suzuki, K. (1999). Group therapy. In M. Kawano (Ed.), *Psychiatric nursing* (pp. 238–245). Tokyo, Japan: Shinko Trading (in Japanese).

Tachimori, H., & Ito, H. (1999). Reliability and validity of the Japanese version of Client Satisfaction Questionnaire. *Seishin Igaku*, *41*, 711–717 (in Japanese).

Tanimoto, C., & Kitaoka-Higashiguchi, K. (2005). Current involvement of public health nurses in patient support groups. *Journal of Japan Academy of Nursing Science*, *25*(3), 102–108 (in Japanese). doi: 10.5630/jans1981.25.3_102

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., … de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60,* 34–42. doi: 10.1016/j.jclinepi.2006.03.012

Toseland, R. W., & Rivas, R. F. (2009). *An introduction to group work practice* (6th ed.). Boston: Allyn & Bacon.

Yalom, I. D. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York, NY: Basic Books.

Yalom, I. D., & Leszcz, M. (2005). *The theory and practice of group psychotherapy* (5th ed.). New York, NY: Basic Books.

Table1. Demographics of subjects

n=248

		n	Mean±S.D. n(%)
Subjects themselves			
Age (yrs)		245	65.4±7.5
	42-59		37(23.1%)
	60-69		75(46.9%)
	70-82		48(30.0%)
Sex	Male	246	77(31.3%)
	Female		169(68.7%)
Type of member	Guest member	248	137(55.2%)
	Host member		111(44.8%)
SHO membership	Non-member	246	52(21.1%)
	Member		194(78.9%)
Persons with mental disorders			
Age (yrs)		246	37.3±10.0
	14-29		37(22.7%)
	30-39		65(39.9%)
	40-49		44(27.0%)
	50-80		17(10.4%)
Sex	Male	247	150(60.7%)
	Female		97(39.3%)
Relation	Child	244	225(92.2%)
	Others		19(7.8%)
Diagnosis	Schizophrenia	246	212(86.2%)
	Others		34(13.8%)
Rehabilitation	Under rehabilitation	245	116(47.4%)
	No rehabilitation		129(52.6%)

・Numbers in the table do not include missing data.

・SHO: Self-help organization

Table 2. Distribution of each item score												n=248				n=174
Factor name		Response alternatives							TFI-19J				TFI-S			
Item No.	Item statement	1	2	3	4	5	6	7	Mean	S.D	Mean -S.D.	Mean +S.D.	Mean	S.D	Mean -S.D.	Mean +S.D.
		n, %														
Instillation of hope (IH)									5.67	1.12						
2	Things seem more hopeful since joining group.		6	14	34	44	60	90	5.65	1.37	4.28	7.01	5.20	1.54	3.66	6.74
			2.4	5.7	13.7	17.7	24.2	36.3								
8	Group helps me feel more positive about my future.	1	4	11	34	62	59	77	5.57	1.30	4.27	6.87	5.03	1.59	3.44	6.62
		0.4	1.6	4.4	13.7	25.0	23.8	31.1								
13	This group inspires me about the future.	1	4	5	30	68	63	77	5.65	1.22	4.43	6.87	4.85	1.63	3.22	6.48
		0.4	1.6	2.0	12.1	27.4	25.4	31.1								
16	This group helps empower me to make a difference in my own life.		4	4	19	67	62	92	5.83	1.15	4.68	6.99	5.06	1.40	3.66	6.46
			1.6	1.6	7.7	27.0	25.0	37.1								
Secure emotional expression (SEE)									5.61	1.01						
3	I feel a sense of belonging in this group.		8	8	30	33	49	120	5.88	1.38	4.50	7.26	4.99	1.42	3.57	6.41
			3.2	3.2	12.1	13.3	19.8	48.4								
5	It's okay for me to be angry in group.	15	16	18	35	52	57	55	4.95	1.76	3.19	6.71	5.02	1.62	3.40	6.64
		6.1	6.5	7.3	14.1	21.0	23.0	22.2								
9	It touches me that people in group are caring toward each other.			7	19	30	62	130	6.17	1.09	5.08	7.25	4.85	1.63	3.22	6.48
				2.8	7.7	12.1	25.0	52.4								
11	In group, the members are more alike than different from each other.	2	7	13	46	68	74	38	5.20	1.29	3.91	6.48	4.97	1.36	3.61	6.33
		0.8	2.8	5.2	18.6	27.4	29.8	15.3								
14	Even though we have differences, our group feels secure to me.			5	16	57	61	109	6.02	1.05	4.97	7.08	4.94	1.51	3.43	6.45
				2.0	6.5	23.0	24.6	44.0								
17	I get to vent my feelings in group.	1	1	7	29	56	63	91	5.79	1.20	4.59	6.99	4.61	1.60	3.01	6.21
		0.4	0.4	2.8	11.7	22.6	25.4	36.7								
19	I can "let it all out" in my group.	4	10	14	44	52	66	58	5.26	1.47	3.79	6.73	4.28	1.82	2.46	6.10
		1.6	4.0	5.7	17.7	21.0	26.6	23.4								
Awareness of relational impact (ARI)									5.62	1.04						
4	I find myself thinking about my family a surprising amount in group.		6	13	30	51	59	89	5.66	1.34	4.32	7.00	5.29	1.50	3.79	6.79
			2.4	5.2	12.1	20.6	23.8	35.9								
6	In group I've really seen the social impact my family has had on my life.	1	6	12	39	49	65	76	5.53	1.36	4.17	6.89	5.57	1.36	4.21	6.93
		0.4	2.4	4.8	15.7	19.8	26.2	30.7								
12	It's surprising, but despite needing support from my group, I've also learned to be more self-sufficient.		9	7	46	54	64	68	5.46	1.33	4.12	6.79	4.64	1.51	3.13	6.15
			3.6	2.8	18.6	21.8	25.8	27.4								
15	By getting honest feedback from members and facilitators, I've learned a lot about my impact on other people.	1	2	16	30	69	58	72	5.52	1.28	4.24	6.81	5.22	1.40	3.82	6.62
		0.4	0.8	6.5	12.1	27.8	23.4	29.0								
18	Group has shown me the importance of other people in my life.			3	27	52	67	99	5.94	1.07	4.86	7.01	5.35	1.40	3.95	6.75
				1.2	10.9	21.0	27.0	39.9								
Social learning (SL)									5.69	1.05						
1	Because I've got a lot in common with other group members, I'm starting to think that I may have something in common with people outside group too.	1	3	12	31	41	48	112	5.82	1.35	4.47	7.17	4.76	1.60	3.16	6.36
		0.4	1.2	4.8	12.5	16.5	19.4	45.2								
7	My group is kind of like a little piece of the larger world I live in: I see the same patterns, and working them out in group helps me work them out in my outside life.		3	10	35	52	68	80	5.66	1.24	4.42	6.90	4.80	1.54	3.26	6.34
			1.2	4.0	14.1	21.0	27.4	32.3								
10	In group sometimes I learn by watching and later imitating what happens.	2	2	9	26	71	72	66	5.59	1.21	4.38	6.80	4.10	1.61	2.49	5.71
		0.8	0.8	3.6	10.5	28.6	29.0	26.6								
Underlined numinals indicate Mean+S.D. > highest possible score																

Table3. Spearman correlation coefficients between factors

	IH	**SEE**	**ARI**	**SL**
SEE	0.802			
ARI	0.851	0.834		
SL	0.803	0.781	0.835	
ALL	0.923	0.936	0.949	0.896

Table4. Confirmatory factor analysis

	χ^2(Df, p)	GFI	RMSEA
Japanese version	429.40 (147, <0.0001)	0.85	0.088(0.079-0.098)
Original	394.29 (143, 0.0001)	0.88	0.077(0.068-0.086)

Table5. Multitrait scaling analysis						
No.	Item statement	Item-scale correlation				Scaling successes (%)
		IH	SEE	ARI	SL	
Instillation of hope (IH)						
2	Things seem more hopeful since joining group.	0.7584	0.6730	0.7386	0.7535	100.0
8	Group helps me feel more positive about my future.	0.7986	0.7277	0.7706	0.7549	
13	This group inspires me about the future.	0.8516	0.7356	0.7724	0.7002	
16	This group helps empower me to make a difference in my own life.	0.7718	0.7443	0.7484	0.7013	
Secure emotional expression (SEE)						
3	I feel a sense of belonging in this group.	0.7556	0.7263	0.7433	0.7153	95.5
5	It's okay for me to be angry in group.	0.4915	0.5704	0.5062	0.4995	
9	It touches me that people in group are caring toward each other.	0.6008	0.6305	0.6070	0.5918	
11	In group, the members are more alike than different from each other.	0.5811	0.5954	0.6307	0.6645	
14	Even though we have differences, our group feels secure to me.	0.6951	0.7270	0.7001	0.6317	
17	I get to vent my feelings in group.	0.6673	0.7580	0.7133	0.6446	
19	I can "let it all out" in my group.	0.6173	0.7094	0.6371	0.6036	
Awareness of relational impact (ARI)						
4	I find myself thinking about my family a surprising amount in group.	0.6516	0.5990	0.6268	0.5946	60.0
6	In group I've really seen the social impact my family has had on my life.	0.6528	0.6865	0.7110	0.6961	
12	It's surprising, but despite needing support from my group, I've also learned to be more self-sufficient.	0.7382	0.7242	0.7165	0.7169	
15	By getting honest feedback from members and facilitators, I've learned a lot about my impact on other people.	0.7062	0.6641	0.7306	0.6875	
18	Group has shown me the importance of other people in my life.	0.7431	0.7379	0.7255	0.7253	
Social learning (SL)						
1	Because I've got a lot in common with other group members, I'm starting to think that I may have something in common with people outside group too.	0.6157	0.6515	0.6461	0.5459	0.0
7	My group is kind of like a little piece of the larger world I live in: I see the same patterns, and working them out in group helps me work them out in my outside life.	0.7329	0.6629	0.7651	0.6464	
10	In group sometimes I learn by watching and later imitating what happens.	0.6959	0.6761	0.6673	0.6135	
N=248						
・Numerals in the IH, SEE, ARI, and SL rows are item-scale Pearson's correlation (corrected for overlap)						
・Underlined correlation coefficients mean scaling failures						
・Scaling successes are percentage of cases in which an item correlates higher with its own scale (corrected for overlap) than with other scales						

Table 6. Concurrent validity: Relations between factors and other scales

	Group Benefit Scale (all subjects, n=239)		CSQ-8 (only guest members, n=148)	
	F	p	F	p
IH	245.6	<0.0001	74.3	<0.0001
SEE	237.2	<0.0001	93.7	<0.0001
ARI	224.3	<0.0001	56.5	<0.0001
SL	135.7	<0.0001	45.1	<0.0001
ALL	302.8	<0.0001	87.9	<0.0001

・Generalized linear mixed model with the group as a random effect

Table7. Concurrent validity :Relations between factors and level of interaction with other members

	Level of interaction with other members							
	Low level n=52		Moderate level n=89		High level n=105		Main effect n=246	
	Mean	SD	Mean	SD	Mean	SD	F	p
IH	5.07	1.24	5.65	1.14	5.98	0.90	11.75	<0.0001
SEE	5.00	1.08	5.63	1.02	5.88	0.84	13.10	<0.0001
ARI	4.97	1.11	5.55	1.05	5.98	0.83	14.37	<0.0001
SL	5.08	1.15	5.66	1.03	6.01	0.87	11.14	<0.0001
ALL	5.02	1.06	5.62	0.98	5.95	0.78	14.19	<0.0001

・Low level: knowing the members only in the program

・Moderate level: knowing some members before the program

・High level: meeting frequently with some members before the program

・Generalized linear mixed model with the group as a random effect

・ALL means an average of all 19

items